CLAIMS

1.   (currently amended) A process for real time analysis of <u>any of</u> text and[[/or]] media content and relating information to the content, comprising the

5   steps of:

analyzing said content in real time;

wherein said analyzing step analyzes said content for semantic and conceptual use;

providing a set of ~~informational~~ <u>reference</u> documents;

10   wherein said ~~informational~~ <u>reference</u> documents comprise any of text, Web, and media documents;

providing a pre-processed analysis of said ~~informational~~ <u>reference</u> documents;

wherein said pre-processed analysis is an analysis of said ~~informational~~

15   <u>reference</u> documents for semantic and conceptual use;

identifying ~~informational~~ <u>reference</u> documents related to said analyzed content using said pre-processed analysis;

providing a user with a description of each identified ~~informational~~ <u>reference</u> document;

20   accepting user input for selecting an identified ~~informational~~ <u>reference</u> document; and

displaying the selected identified ~~informational~~ <u>reference</u> document to the user.

25   2.   (currently amended) The process of Claim 1, wherein said identifying step identifies related ~~informational~~ <u>reference</u> documents by finding ~~informational~~ <u>reference</u> documents that are similar in words, semantically or conceptually, to the analyzed content.

30   3.   (currently amended) The process of Claim 1, further comprising the step of:

storing descriptors for each ~~informational~~ reference document.

retrieving descriptions of each identified ~~informational~~ reference document from said stored descriptors.

5   4.   (currently amended) The process of Claim 1, wherein said set of ~~informational~~ reference documents are stored in a central storage device.

5.   (currently amended) The process of Claim 1, wherein said pre-processed analysis creates a list of words and calculates the frequency that the words

10   appear in said set of ~~informational~~ reference documents.

6.   (original) The process of Claim 5, wherein said pre-processed analysis translates similar words into the same word.

15   7.   (currently amended) The process of Claim 1, wherein said pre-processed analysis generates collocations of words that appear together and calculates the frequency of pairs of words and the frequency of the words appearing together in said ~~informational~~ reference documents.

20   8.   (original) The process of Claim 7, wherein said pre-processed analysis finds relations between collocations to learn their meaning/context.

9.   (currently amended) The process of Claim 1, wherein said pre-processed analysis uses a signature algorithm to calculate signatures for blocks of text,

25   wherein a signature is a vector of words and their weighting within a[[n]] ~~informational~~ reference document; wherein the weighting is determined by the importance of a word in the collocations and within the document.

10.   (original) The process of Claim 9, wherein said pre-processed analysis

30   calculates signatures for Web pages, text tags associated with images, and blocks of text.

11.   (currently amended) The process of Claim 9, wherein said pre-processed analysis creates an index for each word from a signature vector for a[[n]] ~~informational~~ reference document and saves the index, word, text document, and

5   weight of the word into a database that is used to find text documents that have similar signatures.

12.   (original) The process of Claim 9, wherein said pre-processed analysis uses the signatures and weights of the words to create sets of documents that

10   have similar signatures.

13.   (currently amended) The process of Claim 1, further comprising the step of:

       collecting text documents and multimedia from Web pages across the

15   Internet using a Web crawler and placing them into said set of ~~informational~~ reference documents.

14.   (currently amended) A process for real time analysis of <u>any of</u> text and[[/or]] media content in a workflow application and relating information to the

20   content, comprising the steps of:

       automatically analyzing said content in real time as said content is being entered or reviewed by a user;

       wherein said analyzing step analyzes said content for semantic and conceptual use;

25   providing a set of ~~informational~~ <u>reference</u> documents;

       wherein said ~~informational~~ <u>reference</u> documents comprise any of text, Web, and media documents;

       providing a pre-processed analysis of said ~~informational~~ <u>reference</u> documents;

30   wherein said pre-processed analysis is an analysis of said ~~informational~~ reference documents for semantic and conceptual use;

identifying ~~informational~~ reference documents related to said analyzed content using said pre-processed analysis;

wherein said identifying step identifies related ~~informational~~ reference documents by finding informational documents that are similar in words,

5   semantically or conceptually, to the analyzed content;

providing a user with a description of each identified ~~informational~~ reference document;

accepting user input for selecting an identified ~~informational~~ reference document; and

10   displaying the selected identified ~~informational~~ reference document to the user.

15.   (currently amended) A process for real time analysis of media content and relating information to the content, comprising the steps of:

15   extracting metadata from said media content in real time as said content is being viewed by a user;

providing a set of ~~informational~~ reference documents;

wherein said ~~informational~~ reference documents comprise any of text, Web, and media documents;

20   providing a pre-processed analysis of said ~~informational~~ reference documents;

wherein said pre-processed analysis is an analysis of said ~~informational~~ reference documents for semantic and conceptual use;

identifying ~~informational~~ reference documents related to said metadata

25   using said pre-processed analysis;

wherein said identifying step identifies related ~~informational~~ reference documents by finding ~~informational~~ reference documents that are similar in words, semantically or conceptually, to said metadata;

providing a user with a description of each identified ~~informational~~

30   reference document;

accepting user input for selecting an identified ~~informational~~ <u>reference</u> document; and

displaying the selected identified ~~informational~~ <u>reference</u> document to the user.

5

16.   (currently amended) The process of Claim 15, wherein a broadcaster provides customized ~~informational~~ <u>reference</u> documents and specifies their relevance to be used by said identifying step.

10   17.   (currently amended) The process of Claim 15, wherein a producer of said media content provides customized ~~informational~~ <u>reference</u> documents and specifies their relevance to be used by said identifying step.

18.   (original) The process of Claim 15, wherein said extracting step creates
15   metadata for said media content by analyzing said media content if said media content does not have associated in-band metadata.

19.   (currently amendedl) An apparatus for real time analysis of <u>any of</u> text and[[/or]] media content and relating information to the content, comprising:
20   a module for analyzing said content in real time;

wherein said analyzing module analyzes said content for semantic and conceptual use;

a set of ~~informational~~ <u>reference</u> documents;

wherein said ~~informational~~ <u>reference</u> documents comprise any of text,
25   Web, and media documents;

a pre-processed analysis of said ~~informational~~ <u>reference</u> documents;

wherein said pre-processed analysis is an analysis of said ~~informational~~ <u>reference</u> documents for semantic and conceptual use;

a module for identifying ~~informational~~ <u>reference</u> documents related to said
30   analyzed content using said pre-processed analysis;

a module for providing a user with a description of each identified ~~informational~~ reference document;

a module for accepting user input for selecting an identified ~~informational~~ reference document; and

5   a module for displaying the selected identified ~~informational~~ reference document to the user.

20.   (currently amended) The apparatus of Claim 19, wherein said identifying module identifies related ~~informational~~ reference documents by finding

10   informational documents that are similar in words, semantically or conceptually, to the analyzed content.

21.   (currently amended) The apparatus of Claim 19, further comprising:

a module for storing descriptors for each ~~informational~~ reference

15   document.

a module for retrieving descriptions of each identified ~~informational~~ reference document from said stored descriptors.

22.   (currently amended) The apparatus of Claim 19, wherein said set of

20   ~~informational~~ reference documents are stored in a central storage device.

23.   (currently amended) The apparatus of Claim 19, wherein said pre-processed analysis creates a list of words and calculates the frequency that the words appear in said set of ~~informational~~ reference documents.

25

24.   (original) The apparatus of Claim 23, wherein said pre-processed analysis translates similar words into the same word.

25.   (currently amended) The apparatus of Claim 19, wherein said pre-

30   processed analysis generates collocations of words that appear together and

calculates the frequency of pairs of words and the frequency of the words appearing together in said ~~informational~~ reference documents.

26. (original) The apparatus of Claim 25, wherein said pre-processed analysis

5 finds relations between collocations to learn their meaning/context.

27. (currently amended) The apparatus of Claim 19, wherein said pre-processed analysis uses a signature algorithm to calculate signatures for blocks of text, wherein a signature is a vector of words and their weighting within a[[n]]

10 ~~informational~~ reference document; wherein the weighting is determined by the importance of a word in the collocations and within the document.

28. (original) The apparatus of Claim 27, wherein said pre-processed analysis calculates signatures for Web pages, text tags associated with images, and

15 blocks of text.

29. (original) The apparatus of Claim 27, wherein said pre-processed analysis creates an index for each word from a signature vector for an informational document and saves the index, word, text document, and weight of the word into

20 a database that is used to find text documents that have similar signatures.

30. (original) The apparatus of Claim 27, wherein said pre-processed analysis uses the signatures and weights of the words to create sets of documents that have similar signatures.

25

31. (currently amended) The apparatus of Claim 19, further comprising:

a module for collecting text documents and multimedia from Web pages across the Internet using a Web crawler and placing them into said set of ~~informational~~ reference documents.

30

32.   (currently amended) An apparatus for real time analysis of any of text and/[[or]] media content in a workflow application and relating information to the content, comprising:

a module for automatically analyzing said content in real time as said
5   content is being entered or reviewed by a user;

wherein said analyzing module analyzes said content for semantic and conceptual use;

a set of informational reference documents;

wherein said informational reference documents comprise any of text,
10   Web, and media documents;

a pre-processed analysis of said informational reference documents;

wherein said pre-processed analysis is an analysis of said informational reference documents for semantic and conceptual use;

a module for identifying informational reference documents related to said
15   analyzed content using said pre-processed analysis;

wherein said identifying module identifies related informational reference documents by finding informational reference documents that are similar in words, semantically or conceptually, to the analyzed content;

a module for providing a user with a description of each identified
20   informational reference document;

a module for accepting user input for selecting an identified informational reference document; and

a module for displaying the selected identified informational reference document to the user.

25

33.   (currently amended) An apparatus for real time analysis of media content and relating information to the content, comprising:

a module for extracting metadata from said media content in real time as said content is being viewed by a user;

30   a set of informational reference documents;

wherein said ~~informational~~ reference documents comprise any of text, Web, and media documents;

a pre-processed analysis of said ~~informational~~ reference documents;

wherein said pre-processed analysis is an analysis of said ~~informational~~

5   reference documents for semantic and conceptual use;

a module for identifying ~~informational~~ reference documents related to said metadata using said pre-processed analysis;

wherein said identifying module identifies related ~~informational~~ reference documents by finding ~~informational~~ reference documents that are similar in

10   words, semantically or conceptually, to said metadata;

a module for providing a user with a description of each identified ~~informational~~ reference document;

a module for accepting user input for selecting an identified ~~informational~~ reference document; and

15   a module for displaying the selected identified informational document to the user.

34.   (original) The apparatus of Claim 33, wherein a broadcaster provides customized informational documents and specifies their relevance to be used by

20   said identifying step.

35.   (currently amended) The apparatus of Claim 33, wherein a producer of said media content provides customized ~~informational~~ reference documents and specifies their relevance to be used by said identifying step.

25

36.   (original) The apparatus of Claim 33, wherein said extracting step creates metadata for said media content by analyzing said media content if said media content does not have associated in-band metadata.